

EJERCICIOS CON ORDENADOR II.A (ejercicios II.2 y II.3), II.C (ejercicio II.6)

Utilizando ahora un programa informático (R y el paquete R Commander), estimar el intercepto, la pendiente, la bondad del ajuste y la gráfica de la recta de regresión, junto con la nube de puntos, para los datos dados en los ejercicios II.2 y II.3.

SOLUCIÓN

Ejercicio II.2 (viga, precio-resistencia: estimación)

Introducimos los datos del ejercicio I.2, poniendo nombre al fichero, por ejemplo “viga_precio_res”, editando las dos columnas y las cuatro filas:

Datos->Nuevo conjunto de datos->etc

y se marca “aceptar”, para que sean esos los datos activos.

Conviene verificar los datos: “visualizar conjunto de datos”

Regresión, bondad del ajuste

Una vez hemos creado el fichero de los datos:

Estadísticos->ajuste de modelos-> regresión lineal,
aclarando la variable explicativa y explicada.

La salida es, señalando en rojo los datos pedidos:

Call:

```
lm(formula = V2 ~ V1, data = viga_precio_res)
```

Residuals:

	1	2	3	4
	-0.21576	0.57246	-0.44963	0.09293

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.3981	0.4826	2.897	0.1014
V1	1.0147	0.1134	8.948	0.0123 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5409 on 2 degrees of freedom

Multiple R-squared: 0.9756, Adjusted R-squared: 0.9634

F-statistic: 80.07 on 1 and 2 DF, p-value: 0.01226

Los resultados son los mismos que obtuvimos con “lápiz y papel” (salvo posibles errores de redondeo).

Gráfica

Gráficas-> diagrama de dispersión -> en “datos” (elegir las variables x. y), en “opciones”(marcar línea de mínimos cuadrados)

EJERCICIO CON ORDENADOR II.C

Ejercicio II.6 (viga, precio-resistencia: inferencia): error estándar de la regres., desviación estándar de la pendiente, t-valor, ¿el precio depende realmente de la resistencia de la viga? (ie, ¿la resistencia explica el precio?).

SOLUCIÓN

Ya han aparecido los datos en la salida del ejercicio II.A, sólo señalamos en rojo los datos pedidos:

Call:

```
lm(formula = V2 ~ V1, data = viga_ precio_ res)
```

Residuals:

1	2	3	4
-0.21576	0.57246	-0.44963	0.09293

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.3981	0.4826	2.897	0.1014

V1	1.0147	0.1134	8.948	0.0123 *
----	--------	--------	-------	----------

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5409 on 2 degrees of freedom

Multiple R-squared: 0.9756, Adjusted R-squared: 0.9634

F-statistic: 80.07 on 1 and 2 DF, p-value: 0.01226

Mismos valores , salvo errores de redondeo, que los calculados “con lápiz y papel”.

EJERCICIO II.3 (ingreso-empleados)

a) log ingreso-empleados: variables x, z=log y

Datos->Modificar variables conjunto de datos activo->

nombre de la nueva variable: z

Expresión a calcular: log(ingreso) (se escribe log() y se hace doble clic sobre ingreso)

lm(formula = V2 ~ V1, data = log_ingreso_empleados)

Residuals:

1	2	3	4	5	6
0.022333	-0.052067	0.008533	0.021133	0.028733	-0.028667

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.088231	0.042501	-25.61	0.000013819 ***
empleados	0.126452	0.001766	71.60	0.000000228 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.0394 on 4 degrees of freedom

Multiple R-squared: 0.9992, Adjusted R-squared: 0.999

F-statistic: 5127 on 1 and 4 DF, p-value: 2.28 e-07

Mismos resultados que los obtenidos con “lápiz y papel”. Aparecen, además, resultados que tienen que ver con inferencia: errores estándares de la reg., del intercepto y de la pendiente, sus t-valores y sus p-valores: por ejemplo, el p-valor de la pendiente en a) (curva exponencial) es extremadamente pequeño (2.21e-07), lo que indica *claramente*, que rechazamos *con mucha seguridad*

H0: $\beta = 0$,

es decir, que las variaciones en los beneficios son *muy bien explicadas* por las variaciones en el número de empleados.

b) ingreso-empleados: variables x, y, sin aplicar el log

Call:

```
lm(formula = V2 ~ V1, data = ingreso_empleados)
```

Residuals:

1	2	3	4	5	6
3.9248	-0.1458	-2.9964	-4.1170	-1.8175	5.1519

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-12.6659	4.8542	-2.609	0.05947 .
empleados	0.9961	0.2017	4.938	0.00783 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.219 on 4 degrees of freedom

Multiple R-squared: 0.8591, Adjusted R-squared: 0.8239

F-statistic: 24.39 on 1 and 4 DF, p-value: 0.007825

Vemos que la bondad del ajuste no es tan buena como en el caso de la función exponencial.

Aquí el p-valor de la pendiente sigue siendo pequeño (aunque no tanto): 0,00783, es decir todavía no llega al 1 por ciento.

Gáficas: como antes